15. Closeness Counts in Social Choice

Tommi Meskanen
Department of Mathematics, University of Turku, Finland

Hannu Nurmi
Department of Political Science, University of Turku, Finland

1. Introduction

A working democratic system of government is based on voting and elections. A glance at the literature reveals an astonishing variety of systems used for electing persons to political offices, for choosing policy alternatives and for enacting legislation. While it is undoubtedly true that voting is just a necessary condition for democracy, it is remarkable how many and how different systems are used for apparently the same purpose, viz. to find out the 'will of the people'.

In early 1980's Riker called attention to the fact that not only are the voting systems apparently different, but also the voting outcomes may vary a great deal depending on the voting system used (Riker 1982). In extreme cases, any alternative under consideration can be rendered the winner by applying a suitable procedure without changing anyone's opinion about the alternatives. Although extreme and theoretical, these examples show that the choice of voting system is an important aspect of democratic rule.

All voting systems seem, however, to satisfy the minimal requirement of democratic choice, viz. that the outcome depends on votes cast. So, why all this variety? This article dwells on this question. We shall first trace the historical development of the voting procedures by giving a brief tour through recent history mentioning the authors who have brought to discussion new voting systems. Thereafter, we argue that most systems can be characterized by a goal state, i.e. consensus, and that each procedure measures the distance from this consensus by a metric.

2. A Sketch of the History of Voting Systems

Historical record on the invention of the first voting methods does not exist, but a version of plurality voting was commonly used in one of the city

states of ancient Greece, Sparta. Aleskerov (1999: 1) cites the account of the Greek historian Plutarch on the election of Council of Elders: 'Whoever [of the candidates] was greeted with the most and loudest shouting, him they declared elected'.

Sparta was probably not the first place where this method was used, but earlier reports on its use have not come our way. The next record emerges from the first century A.D. in a letter written by Pliny the Younger to Titus Aristo. This letter was brought to wider scholarly attention by R. Farquharson (1969). It reveals that the Roman senate resorted to several voting systems, plurality voting being one of them. The letter hints that also those voting systems that are today used in legislative decision making in parliaments, notably the amendment and successive procedures, were known in the Roman senate.

The first systematic treatises that include voting system proposals appeared in the late 18th century. In the middle ages there were a few scattered accounts suggesting voting rules (McLean and Urken 1995), but these did not typically include systematic comparisons with existing systems, let alone theoretical discussion on their properties. But in 1770 Jean-Charles de Borda presented to the Royal French Academy a strong criticism of the prevailing plurality voting system augmented with a proposal for a new system, currently known as the Borda count (BC, for brevity) (McLean and Urken 1995: 81–89). Given a set of individual complete and transitive preference relations (rankings) over the alternatives, BC works as follows. Each voter gives $k-1$ points to his/her (hereafter his) first ranked alternative, $k-2$ to the second ranked one, …, zero points to the last-ranked one. The Borda score of each alternative is the sum of points it has received from all voters. The collective Borda ranking is the order of the Borda scores: the larger the score, the higher the alternative in the collective ranking.

Borda's contemporary, Marquis de Condorcet, proposed, some fifteen years later, several voting systems of which one, today called Condorcet's practical method, seems to have had practical applications. This method calls for each voter to indicate his first and second ranked alternatives. Should one of the alternatives be ranked first by more than 50% of the voters, it is elected. Otherwise, the alternative which has been ranked first or second by more voters than any other alternative is elected. Nanson (1882) calls this Condorcet's practical method and reports that it was used in Geneva. Yet another system, called least reversal method, is associated with the name of Condorcet. Suppose that pairwise majority comparisons yield a cycle. Then one reverses the majority opinion with smallest support and checks if the resulting order contains other cycles. If it does, one reverses the pair with next smallest support and so on. Eventually an acyclic majority relation emerges. This is adopted as the collective decision.

Nearly a hundred years later C. L. Dodgson wrote a series of pamphlets analyzing the voting systems used in colleges of Oxford (McLean and Urken

1995). At least the following methods are mentioned in his works: methods of simple or absolute majority, method of pairwise elimination, method of elimination, method of marks (i.e. cumulative voting or range voting), method of nomination (i.e. successive procedure), method of procedure (i.e. Borda count) and a method today called the Dodgson's method. The expressions in parentheses are terms currently used in discussing the methods in question. Dodgson's own proposal was presented in the pamphlet dated in March 1876.[1] It amounts to electing a Condorcet winner whenever one exists and otherwise looking for the alternative that is closest to a Condorcet winner, closest here meaning that the alternative needs a smaller number of pairwise preference changes of voters than any other alternative.

In contrast to Dodgson his contemporary E. J. Nanson (1882) was well aware of the preceding developments in the theory of voting procedures. He discussed voting systems in three categories:

Class one (single scrutiny): single vote method, double vote method, Borda count.

Class two (voters have to vote more than once): French method of double elections.

Class three (more than one scrutiny, but the voters vote only once (preferentially)): Ware's method, Venetian method (two ballots: on the first one the voters vote for two, in the second for one candidate), Condorcet's practical method(two ballots: on the first each voter has one vote, on the second two votes; no second ballot if a candidate receives an absolute majority on the first ballot), the proposed method, i.e. Nanson's method.

The last mentioned one is a clever Borda elimination method based on the observation that, although a Condorcet winner is not necessarily the Borda winner, the former receives necessarily a higher than average Borda score. Hence, by eliminating alternatives that get an average or smaller Borda score, one can rest assured that an eventual Condorcet winner is not eliminated. By continuing the elimination one finally ends up with the Nanson winner (or a tie). There is a version of Nanson's method in which candidates are eliminated one by one so that at each stage of computation the candidate with the lowest Borda score is deleted. This version is known as Baldwin's method.

After Nanson's article the study of voting procedure became eclipsed by increasing interest in proportional election systems. Of particular interest is the work of Thomas Hare (1865). The system he devised is nowadays called

[1] In fact, it might be slightly unfair to attribute this proposal to Dodgson since he explicitly states that in the absence of a Condorcet winner the outcome should be 'no Election'. However, having said this he proceeds to show that the two ways of making a choice from a majority cycle, the plurality and runoff systems, are erroneous in not electing the candidate that can become the Concorcet winner after smaller number of pairwise preference changes than any other candidate. Hence, calling his proposal Dodgson's method has, after all, a modicum of plausibility.

the Hare system.[2] It is an elimination method whereby one looks for a candidate that is first ranked by more than 50% of the voters. If there is no such candidate, the one whom the smallest number of voters ranks first is eliminated. The elimination is continued until one of the non-eliminated candidates is top-ranked by more than half of the electorate.

The single-winner voting systems were not the topic of a major treatise until 1940's when Duncan Black published some articles on the subject. Black's *magnum opus* was published in 1958. His suggestion for a voting procedure is preferential balloting with the eventual Condorcet winner elected. Otherwise the Borda winner is elected. The method thus combines the decisiveness of the Borda count with the binary winning intuition of Condorcet.

The voting system devised by Kemeny (1959) is explicitly based on the search for the consensus that is closest to the prevailing distribution of opinions. Once individual preference rankings are given, one looks for the (consensus) preference ranking that would result from the given one after the smallest number of binary changes of opinions by individual voters.

Slater (1961) suggests a method that is very similar to Kemeny's. Given a preference profile, one constructs the corresponding tournament matrix, *i.e.* a $0-1$ matrix where 1 in i-th row and j-th column indicates that i-th alternative defeats the j-th one by a majority. Otherwise the (i, j) entry is 0. This is the observed tournament matrix over the alternatives. This may or may not be cyclic. One then compares this matrix with all those tournaments that represent complete and transitive relations over the same alternatives. Finally, one computes the distance of the observed matrix to all those representing complete and transitive rankings. The distance is measured by the number of binary switches of preferences needed to transform the observed matrix into one representing a complete and transitive relation (Nurmi 2002).

In 1970's the approval voting system was introduced by Steven Brams (1976).[3] In this system every voter can give each alternative either one or zero votes. The alternative with the largest vote sum is elected.

Perhaps the first systematic comparison of voting systems was published by Straffin (1980). This study introduces several performance criteria for voting systems. Also the systems attributed to Hare and Coombs are analyzed in the book. Coombs's system works exactly as Hare's except that the candidates with the largest number of last ranks are eliminated.

---

[2] There are several minor variations of the system, the best known being the single transferable vote which is designed for multi-member constituencies. In Australia, the Hare system is known as the Hare-Spence system and in Scandinavia as the Andre system. See Lakeman and Lambert (1955) and Doron (1979). The history of the system is briefly described in Hill (1988).

[3] An account on how several groups of scholars came up with the same idea roughly simultaneously is given in (Brams and Fishburn 1983).

In late 1970's Richelson (1975, 1978*a*, 1978*b*, 1979, 1981) conducted an extensive survey and analytic comparison of voting systems introducing, i.a. reverse plurality, nowadays better known as antiplurality system. In this system each voter's last ranked alternative gets zero votes, while all the others receive one vote from the voter. The alternative with the largest vote sum wins.

The above mentioned are all voting systems where each voter casts a ballot containing either ranking of candidates, a list of approved candidates or the name (or number) of a single (most-preferred) candidate. In his widely cited text Riker (1982) discusses two other systems both of which require utility values of alternatives. Bentham's method defines the winner as the alternative with the largest sum of utility values, while the Nash winner is the alternative with the largest product of utility values. The utility scales are fixed so that 1 is the value of the first ranked and 0.5 the value of the last ranked alternative.[4]

An explicitly consensus oriented system is due to Litvak (1982). Given a preference profile over k alternatives, one first agrees on the numbering of alternatives $1,\ldots,k$. Then each individual preference ranking is transformed into a vector of k components so that component i indicates how many alternatives are placed ahead of the i-th one in the ranking under consideration. Each such vector thus consists of numbers $0,\ldots,k-1$. Litvak's rule looks for the k-component vector V that is closest to the observed preferences in the sense that the sum of component-wise absolute differences between the reported preference vectors and V is minimal. For example, in a 5-person preference profile where three persons have the preference $A \succ B \succ C$ and two the preference $B \succ C \succ A$, the Litvak sums of all six rankings are:

$$A \succ B \succ C : 8 \qquad B \succ A \succ C : 10 \qquad C \succ A \succ B : 20$$
$$A \succ C \succ B : 14 \qquad B \succ C \succ A : 12 \qquad C \succ B \succ A : 16$$

Thus, the Litvak ranking is $A \succ B \succ C$.

Tideman's (1987) method is a sequential one proceeding from the largest pairwise victory of an alternative against another. Assume that x_1 and x_2 form such a pair. Denoting by v_i the number of votes x_i receives, this means that $v_1 - v_2$ is the largest difference of votes. We preserve this pair by drawing an arrow from x_1 to x_2. We then look for the next largest victory margin and preserve it similarly. Proceeding in this manner we may encounter a situation where preserving a pair would create a cycle. In these situations the respective relation between alternatives is ignored, i.e. no arrow is drawn. Otherwise we proceed in decreasing order of victory margins. Eventually all alternatives are located in a directed graph with no cycles.

---

[4] Resorting to the common 0–1 normalization of utilities would effectively make every voter a vetoer.

The root of the graph, that is, the starting point is the winning alternative. Since by construction the cycles are excluded, Tideman's method yields a ranking over the set of alternatives.

In late 1980's I. D. Hill (1988) proposed a variation of Hare's system. In single-winner elections it proceeds exactly as Hare's method, but exempts the eventual Condorcet winner from elimination. In other words, in case none of the remaining candidates is ranked first by more than 50% of the voters, the candidate with the smallest number of first ranked is eliminated, unless this candidate happens to be the Condorcet winner among the remaining candidates. Since the Condorcet winner is not eliminated, it wins.

Schulze's (2003) method, in turn, determines the 'beatpaths' from each alternative x to every other alternative y. The path consists of ordered alternative pairs in the chain $x, w_1 \ldots, w_j, y$ – where $w_1, \ldots, w_j$ are alternatives – that leads from x to y. The strength of the beatpath is the smallest margin of victory in the chain. There are typically several beatpaths leading from one alternative to another. Each of them, thus, has the strength of its weakest link. Denote now by S_{xy} the strength of the strongest path from x to y. In other words, all but the strongest path between any two alternatives are ignored. Schulze's (potential) winner is the alternative x_S for which $S_{x,y} \geq S_{yx,}$ for all alternatives y.[5] Schulze proves that at least one such candidate always exists. Moreover, the relation $S_{xy} > S_{yx}$ is transitive. That is, if $S_{xy} > S_{yx}$ and $S_{yz} > S_{zy},$, then also $S_{xz} > S_{zx},$ for all alternatives x, y, z.

The above list of systems is fairly long, but not exhaustive. The existence of such variety of systems poses the question of why do we have them. Each system is minimally democratic in the sense that the outcome depends, inter alia, on the opinions expressed by the voters. Surely there must then be some more profound motivation for these systems, a motivation that accounts for the plausibility of each system as distinguished from others. In the following we shall approach some of the above systems by looking at the consensus state which they are aimed at. It will be seen that the differences in many systems can be accounted for by differences in the consensus state they are based on and by the method of measuring the distance from that state.

3. Consensus States

Apart from Dodgson's, Slater's and Kemeny's rules the above systems do not explicitly contain an idea that the voting outcome would be such a consensus state which is nearest to the reported preferences. Yet, upon closer scrutiny this idea can be associated with most systems outlined above. To focus on Kemeny, the state from which the distance to the observed profile is

[5] In case there are several potential winners, Schulze (2003) suggest a tie-breaking procedure which, however, will not be discussed here.

measured is one of unanimity regarding all positions of the ranking of alternatives, i.e. the voters are in agreement about which alternative is placed first, which second etc. throughout all positions.

Another possible consensus state is one where unanimity prevails concerning which alternative is the best. For example, the plurality voting system can be viewed in the light of this consensus state. The observed preference profile is the further away from the consensus, the more opinions have to be changed in order to make one alternative unanimously first ranked.

In the context of Dodgson's method, the consensus is reached when there is a Condorcet winner in the profile. If there is one in the observed profile, we are already there, but in case there isn't, one looks for the number of opinion changes needed to make each alternative the Condorcet winner.

Another possible idea of consensus is one where there is a Condorcet ordering. There exists a Condorcet ordering over alternatives if for every subset of alternatives there is a Condorcet winner. In other words, there is an alternative, say x_1, that is a Condorcet winner in the entire candidate set X (and consequently in every subset of X to which x_1 belongs), there is Condorcet winner, say x_2, in $X \setminus x_1$, and so on.

It turns out that with these interpretations of the consensus state and the associated distance measures we can characterize many of the above voting systems. Obviously the three different goal consensus states are not adequate to explain the variety of procedures. Hence, we focus next to the distance measures.

4. Metrics

We define a distance function over a set P (of points, for example) as any function $d: P \times P \to R^+$, where R^+ is the set of non-negative real numbers. A distance function d_m is called a metric if the following conditions are met for all elements x, y, z of P:

1. $d_m(x,x) = 0$,
2. if $x \neq y$, then $d_m(x,y) > 0$,
3. $d_m(x,y) = d_m(y,x)$, and
4. $d_m(x,z) \leq d_m(x,y) + d_m(y,z)$.

Substituting preference relations for elements in the above conditions we can extend the concept of distance function to preference relations. But these conditions leave open the way in which numerical values consistent with the above conditions are assigned to distances between two relations. Kemeny's proposal is the following (Baigent 1987*a*,*b*; Kemeny 1959).

Let R and R' be two rankings. Then their distance is:

$$d_K(R,R') = |\{(x,y) \in X^2 \mid R(x) > R(y), R'(y) > R'(x)\}|.$$

Here we denote by $R(x)$ the number of alternatives worse than x in a ranking R. This is called inversion metric.

The distance of two rankings is the number of inversions of consecutive choices needed to transform one ranking into another. This can be easily seen because:

- If two consecutive options are in different order they must be inverted at some point
- If two consecutive options are in an identical order they need not to be inverted.

The distance between two preference profiles of the same size, i.e. involving equal number of voters and the same alternatives, is, then, the sum of the distances between the individual rankings. That is,

$$d_K(P,P') = \sum_{i=1}^{n} d_K(R_i, R_i')$$

where the profile P consists of rankings $R_1,\ldots,R_n$ and the profile P' of rankings $R_1',\ldots,R_n'$. Similarly, we can measure the distance between a profile P and a set of profiles $\mathbf{S}$,

$$d_K(P,\mathbf{S}) = \min_{P' \in \mathbf{S}} d_K(P,P').$$

We can generalize any distance function of rankings this way.

Let $U(R)$ denote an unanimous profile where every voter's ranking is R. Kemeny's rule results in the ranking $\overline{R}$ so that

$$d_K(P,U(\overline{R})) \leq d_K(P,U(R)) \quad \forall R \in \mathscr{R} \setminus \overline{R}$$

where P is the observed profile and $\mathscr{R}$ denotes the set of all possible rankings. If all the inequalities above are strict then $\overline{R}$ is the only winner.

The Kemeny winner profile is sometimes called the Kemeny median. Barthélemy and Monjardet (1981) show that this is, indeed, a natural way to look at Kemeny's rule. Let R_1 and R_2 be two binary relations over the same set of alternatives so that $R_1 \subseteq R_2$. In other words, all alternative pairs that are in relation R_1 with each other, are also in relation R_2 with each other, but not necessarily the other way around. Denote by $[R_1,R_2]$ the interval between R_1 and R_2, i.e. the set of all relations S such that $R_1 \subseteq S \subseteq R_2$.

Let now i range from individual 1 to individual n and $(x,y) \in R_i$ denote i's preference of x over y. We define R_m as follows: $(x,y) \in R_m$ if and only if $(x,y) \in R_i$ for at least $[(n+2)/2]$ individuals. Similarly, let R_e be defined so that $(x,y) \in R_e$ if and only if $(x,y) \in R_i$ for at least $[(n+1)/2]$ individuals. Here $[x]$denotes the integer part of the number x. It can be seen immedi-

ately that $R_m \subseteq R_e$. Barthélemy and Monjardet define the median interval in the profile as the interval $[R_m, R_e]$ and a relation is called a median relation if and only if it is in the median interval. They then show that a relation is a median relation in a profile if and only if it minimizes the distance between itself and the the profile when the distance (or in Barthélemy and Monjardet's terminology, remoteness) is measured with d_K. So, the Kemeny winning relation is in a natural sense the median relation.

To turn now to the Borda count, let us consider an observed profile P. For a candidate x we denote by $\mathbf{W}(x)$ the set of all profiles where x is first-ranked in every voter's ranking. Clearly in all these profiles x gets the maximum Borda points. We consider these as the consensus states for the Borda count (Farkas and Nitzan 1979; Nitzan 1981).

For a candidate x, the number of alternatives above it in any ranking of P equals the number of points deducted from the maximum points. This is also the number of inversions needed to get x in the winning position in every ranking. Thus, using the metric above, w_B is the Borda winner if

$$d_K(P, \mathbf{W}(w_B)) \leq d_K(P, \mathbf{W}(x)) \quad \forall x \in X \setminus w_B.$$

The plurality system is obviously also directed at the same consensus state as the Borda count, but its metric is different. Rather than counting the number of pairwise preference changes needed to make a given alternative unanimously first ranked, it minimizes the number of individuals having different alternatives ranked first.

To make this intuition precise, we define a discrete metric as follows:

$$d_d(R, R') = \begin{cases} 0 \text{ if } R = R' \\ 1 \text{ otherwise.} \end{cases}$$

As above the distance between two preference profiles of the same size is

$$d_d(P, P') = \sum_{i=1}^{n} d_d(R_i, R_i'),$$

where the profile P consists of rankings $R_1, \ldots, R_n$ and the profile P' of rankings $R_1', \ldots, R_n'$. That is, the distance indicates the number of rankings that differ in the two profiles. Note that we consider the profiles always in such a way that the order of the rankings in the profiles is irrelevant. Also, the distance between a profile P and a set of profiles $\mathbf{S}$ is, similarly,

$$d_d(P, \mathbf{S}) = \min_{P' \in \mathbf{S}} d_d(P, P').$$

The unanimous consensus state in plurality voting is one where all voters have the same alternative ranked first. With the metric, in turn, we tally, for each alternative, the number of voters in the observed profile who do not have this alternative as their first ranked one. The alternative for which this

number is smallest is the plurality winner. The plurality ranking coincides with the order of these numbers.

Using this metric we have for the plurality winner w_p,

$$d_d(P, \mathbf{W}(w_p)) \leq d_d(P, \mathbf{W}(x)) \quad \forall x \in X \setminus w_p.$$

The only difference to the Borda winner is in the metric used.

Dodgson's system is based on a different idea of the goal state, *viz.* one where there is a Condorcet winner. For any candidate x we denote by $\mathbf{C}(x)$ the set of all profiles where x is the Condorcet winner.

Provided that a Condorcet winner exists in the observed profile, the Dodgson winner is this alternative. Otherwise, one constructs, for each alternative x, a profile in $\mathbf{C}(x)$ which is obtained from the observed profile P by moving x up in one or several voters' preference orders so that x emerges as the Condorcet winner. Obviously, any alternative can thus be rendered a Condorcet winner. It is also clear that for each alternative there is a minimum number of such preference changes involving the improvement of x's position *vis-à-vis* other alternatives that are needed to make x the Condorcet winner. The Dodgson winner w_D is then the alternative which is closest, in the sense of Kemeny's metric, of being the Condorcet winner. That is,

$$d_K(P, \mathbf{C}(w_D)) \leq d_K(P, \mathbf{C}(x)) \quad \forall x \in X \setminus w_D.$$

Dodgson's method is thus characterized by Kemeny's inversion metric combined with a goal state where a Condorcet winner exists.

Litvak's procedure results in the ranking that is nearest to the observed one in terms of minimizing the sum (over individuals) of absolute rank position differences of alternatives in the former and the latter. As the position numbers play an important role in the Borda count, one would expect that Litvak's method is similar to the Borda count. This is, however, not the case (Nurmi 2004). For example, Litvak's procedure may end up with a Condorcet loser being ranked first which is never the case under the Borda count.

Litvak's procedure also differs from Kemeny's. To see this, consider the distance of $A \succ B \succ C \succ D$ to $C \succ D \succ A \succ B$, on the one hand, and to $D \succ C \succ B \succ A$, on the other. In Kemeny's sense, the latter difference is larger than the former, while in Litvak's sense they are equidistant from $A \succ B \succ C \succ D$. We define the Litvak metric formally as follows:

$$d_L(R, R') = \sum_{x \in X} |\, R(x) - R'(x) \,|.$$

This metric is sometimes called the Manhattan metric. The Litvak winning ranking R_L has the property

$$d_L(P, U(R_L)) \leq d_L(P, U(R)) \quad \forall R \in \mathscr{R} \setminus R_L$$

exactly like the Kemeny winner except for the different metric.

The Litvak and Kemeny metrics are related in the following way:

- If a ranking R' is derived from ranking R by only moving one candidate up (or down) by n steps then

$$2d_K(R,R') = 2n = d_L(R,R').$$

- If a ranking R cannot be turned into ranking R' by switching adjacent candidates without moving at least one candidate at some point first up then down (or first down then up), then

$$d_K(R,R') < d_L(R,R') < 2d_K(R,R').$$

Let us consider an example: ranking $A \succ B \succ C$ and its reversal $C \succ B \succ A$. The distance of these orders is 3 using Kemeny metric and 4 using Litvak metric: $3 < 4 < 6$.

Thus, because of the first property, we have that for any P and x

$$d_L(P, \mathbf{W}(x)) = 2d_K(P, \mathbf{W}(x))$$

and

$$d_L(P, \mathbf{C}(x)) = 2d_K(P, \mathbf{C}(x)).$$

In words, we find the same Borda count and Dodgson winners using either of the metrics.

For the sake completeness we shortly consider what we get if we combine the discrete metric with consensus states $U(R)$ or $\mathbf{C}(x)$. In the former case simply the most popular ranking is selected; *that is, the ranking with least opposition*. Although nowhere in use within our knowledge, this system would make sense in situations where there are strong dependencies between alternatives so that each ranking is a policy program. In this interpretation, the policy program with most support would seem a plausible winner.

The latter case, i.e. the one where $\mathbf{C}(x)$ is the consensus state, is more interesting. We can formally define the winner w_Y as the option with property

$$d_d(P, \mathbf{C}(w_Y)) \leq d_d(P, \mathbf{C}(x)) \quad \forall x \in X \setminus w_Y.$$

In other words, we find the largest set of voters such that the Condorcet winner exists. This system has been attributed to H. P. Young (Smith 2005).

5. From Rankings to Matrices

Preference profiles provide relatively rich information about individual opinions. In some cases this information is essentially reduced. For instance, the winner may be determined on the basis of pairwise comparisons of alternatives. Sometimes only the winner of each comparison is recognized,

while in other systems also the victory margin plays a role in determining the overall winner.

We denote by V the outranking matrix where entry V_{xy} indicates the number of voters in profile P preferring candidate x to candidate y. The diagonal entries are left blank.

A metric on outranking matrices can now be defined as follows: if V and V' are the outranking matrices of profiles P and P' then

$$d_V(P,P') = \frac{1}{2}\sum_{x,y\in X} |\, V_{xy} - V'_{xy} \,|.$$

In other words, the distance tells us how much pairwise comparisons differ in the corresponding outranking matrices.

This metric is very similar to the inversion metric and, indeed, we find the same Borda count (or, alternatively, Kemeny) winners using this metric instead.

Let us now turn to systems where the goal is a Condorcet winner. As pointed out above, in the Condorcet least-reversal system the winner is the candidate which can be turned into Condorcet winner with minimum number of reversals of pairwise comparisons. That is, the Condorcet least-reversal system winner w_{lr} is the candidate with property

$$d_V(P,\mathbf{C}(w_{lr})) \leq d_V(P,\mathbf{C}(x)) \quad \forall x \in X \setminus w_{lr}.$$

Copeland's procedure has also a similar goal state as Dodgson's and Condorcet's least reversal one, namely, one with a Condorcet winner. Given an observed profile P one considers the corresponding tournament matrix T where entry $T_{xy} = 1$ if majority of the voters in profile P are preferring candidate x to candidate y. Otherwise $T_{xy} = 0$.

Now, the Condorcet winner is seen as a row in T where all $k-1$ non-diagonal entries are ones.

We define yet another, viz. tournament, distance between profiles P and P' with tournament matrices T and T' as

$$d_T(P,P') = \frac{1}{2}\sum_{x,y\in X} |\, T_{xy} - T'_{xy} \,|.$$

The Copeland winner w_C is the alternative that wins the largest number of comparisons with other candidates i.e. has the smallest number of zeros in its row in the tournament matrix. Thus the winner w_C is the candidate that comes closest to win every other candidate, that is, using the distance above,

$$d_T(P,\mathbf{C}(w_C)) \leq d_T(P,\mathbf{C}(x)) \quad \forall x \in X \setminus w_C.$$

Obviously, the goal states of Condorcet least-reversal system and Cope-

land's system are the same, but metrics differ. The latter pays no attention to majority margins, while the former depends on them.[6]

The idea of the Condorcet winner can be extended into a Condorcet ordering. Let P be a profile and X' be a subset of the set of candidates X. We denote by $P|_{X'}$ the profile obtained when we consider only the candidates in X' and dismiss all other candidates. We say that a profile P has a Condorcet ordering if for every $X' \subset X$ the profile $P|_{X'}$ has a Condorcet winner.

The fact that there is a Condorcet ordering in a profile P over X means that there is a candidate, say $w_1 \in X$ that wins every other candidate in pairwise comparisons with a majority of votes. Moreover, there is a candidate, say w_2, that wins every other candidate except w_1 in pairwise comparisons with a majority of votes. In addition, there is a candidate, say w_3, winning every other candidate except w_1 and w_2 in pairwise comparisons with a majority of votes. And so on. The Condorcet ordering is $w_1 \succ w_2 \succ \cdots \succ w_k$, where $k = |X|$.

It can be easily seen that P has a Condorcet ordering if and only if the tournament matrix of P does not have any cycles, i.e. there are no candidates $x_1, x_2, \ldots, x_i$ such that $T_{x_1x_2} = T_{x_2x_3} = \ldots = T_{x_{i-1}x_i} = T_{x_ix_1} = 1$.

We denote by $\mathbf{Co}(R)$ the set of all profiles that have a Condorcet ordering R. This set could be viewed as a natural goal state for a decision rule. But, as above, there are several different methods to determine which ordering is at smallest distance from an observed profile that does not have a Condorcet ordering.

Two candidates for such measures are (i) the number of pairwise comparisons that need to be reversed and (ii) the number of pairwise losses with majority that need to be turned into victories. That is, we could measure the differences in either outranking or tournament matrices. The latter of these methods is connected to the work of Slater (1961).

$$d_T(P, \mathbf{Co}(R_{Sl})) \leq d_T(P, \mathbf{Co}(R)) \quad \forall R \in \mathscr{R} \setminus R_{Sl}.$$

The former method does not have a name:

$$d_V(P, \mathbf{Co}(R_{U_1})) \leq d_V(P, \mathbf{Co}(R)) \quad \forall R \in \mathscr{R} \setminus R_{U_1}.$$

This method is, thus, based on minimizing the distance between the observed profile and one that has a Condorcet ordering, the crucial feature being that the distance is measured in terms of outranking metric. At first sight this method resembles Dodgson's, but differs from the latter in tallying only differences in the outranking matrices.

Also the other three winners we can define with the metrics mentioned above are yet to be named:

---

[6] See, Klamler (2005) for another distance based characterization of the Copeland rule.

$$d_K(P,\mathbf{Co}(R_{U_2})) \le d_K(P,\mathbf{Co}(R)) \quad \forall R \in \mathscr{R} \setminus R_{U_2}$$
$$d_L(P,\mathbf{Co}(R_{U_3})) \le d_L(P,\mathbf{Co}(R)) \quad \forall R \in \mathscr{R} \setminus R_{U_3}$$
$$d_d(P,\mathbf{Co}(R_{U_4})) \le d_d(P,\mathbf{Co}(R)) \quad \forall R \in R \setminus R_{U_4}.$$

6. Elimination Metrics

Above we have defined the metric counting the differences in the individual rankings (d_K, d_L, d_V), the pairwise defeats (d_T), and the nonidentical rankings (d_d). Another approach is to consider the candidates that vary in two preference profiles.

There are several recursive elimination systems that are based on the order of the eliminations of the candidates.

Let R be some ranking of the candidates, $x_1 \prec x_2 \prec \cdots \prec x_k$ and P the voting profile we are considering. We begin by comparing the Hare system and the function

$$F_P(R) = \sum_{i=1}^{k} (n+1-d_d(P_{i-1},\mathbf{W}(x_i)))\,(n+1)^{k-i}$$

where $P_{i-1} = P\,|_{X\setminus\bigcup_{j=1}^{i-1} x_j}$.

As was pointed out above, on the first round of the Hare system we eliminate the candidate with the smallest number of first places. The function F_P gets its smallest value when the function

$$F'_P(x_1) = (n+1-d_d(P,\mathbf{W}(x_1)))\,(n+1)^{k-1}$$

gets its smallest value. This happens when x_1 is the candidate with the smallest number of first places in P and thus $d_d(P,\mathbf{W}(x_1))$ is maximal. Note that the part

$$\sum_{i=2}^{k} (n+1-d_d(P_{i-1},\mathbf{W}(x_i)))\,(n+1)^{k-i}$$

of $F_P(R)$ is always smaller than $(n+1)^{k-1}$. That is why we do not need to care about candidates $x_2, x_3, \ldots, x_k$ at this point.

On the second round of the Hare system we eliminate the candidate with the smallest number of first places after we have removed the candidate that was eliminated at the previous round. The function F_P gets its smallest value when the function

$$F''_P(x_1,x_2) = (n+1-d_d(P,\mathbf{W}(x_1)))\,(n+1)^{k-1} + (n+1-d_d(P|_{X\setminus x_1},\mathbf{W}(x_2)))\,(n+1)^{k-2}$$

gets its smallest value. This happens when x_1 is the candidate with the

smallest number of first places and x_2 is the candidate with the smallest number of first places after x_1 is eliminated from P. Note again that the part

$$\sum_{i=3}^{k} (n+1-d_d(P_{i-1}, \mathbf{W}(x_i)))\,(n+1)^{k-i}$$

of $F_P(R)$ is always smaller than $(n+1)^{k-2}$. That is why we do not need to care about candidates $x_3, x_4, \ldots, x_k$ at this point. If we continue these considerations we find that the function F_P gets its smallest value when $x_1, x_2, \ldots$ is the order of the eliminations using the Hare system.

We now turn this function into a metric. A metric is always symmetric. Let P and P' be two preference profiles of n rankings and their sets of candidates X and X' have k and k' elements, respectively. We write shortly $P_{i-1} = P|_{X \setminus \bigcup_{j=1}^{i-1} x_j}$. Let the metric for the Hare system be

$$d_H(P,P') = \min\left\{ \sum_{i=1}^{m} (n+1-d_d(P_{i-1}, \mathbf{W}(x_i)))\,(n+1)^{k-i} + \right.$$
$$\left. \sum_{i=1}^{m'} (n+1-d_d(P'_{i-1}, \mathbf{W}(x'_i)))\,(n+1)^{k'-i} \mid x_i \in X, x'_i \in X', P_m = P'_{m'} \right\}.$$

Because $d_H(P,P)$ must be zero for any P, we only calculate the distance until $P_m = P'_{m'}$ for some m and m'.

Note that when we calculate the distance between a profile P and the sets $\mathbf{W}(x)$ (or, alternatively, $\mathbf{C}(x)$) there is always a profile in the closest set $\mathbf{W}(x)$ (or $\mathbf{C}(x)$) such that the second part of the distance function is zero.

When we combine this metric with goal states $\mathbf{W}(x)$ we get the Hare system. If we instead use the goal states $\mathbf{C}(x)$ we get a system named after Hill.

For the Coombs method we have the metric

$$d_C(P,P') = \min\left\{ \sum_{i=1}^{m} (d_d(P_{i-1}, \mathbf{L}(x_i)) + 1)\,(n+1)^{k-i} \right.$$
$$\left. + \sum_{i=1}^{m'} (d_d(P'_{i-1}, \mathbf{L}(x'_i)) + 1)\,(n+1)^{k'-i} \mid x_i \in X, x'_i \in X', P_m = P'_{m'} \right\},$$

where we denote by $\mathbf{L}(x)$ the set of all profiles where x is last-ranked in every voter's ranking Finally, for the Baldwin, also called Borda runoff, method we have

$$d_B(P,P') = \min\left\{ \sum_{i=1}^{m} (kn - d_K(P_{i-1}, \mathbf{W}(x_i)))\,(kn)^{k-i} \right.$$
$$\left. + \sum_{i=1}^{m'} (k'n - d_K(P'_i - 1, \mathbf{W}(x'_i)))\,(k'n)^{k'-i} \mid x_i \in X, x' \in X', P_m = P'_{m'} \right\}.$$

7. Two More Systems

Turning now to somewhat more recent systems, Tideman's (1987) procedure operates on pairwise majority margins. At every step the algorithm fixes one comparison between two candidates such that a minimum amount of voters are disappointed. The result is a complete directed graph without cycles *i.e.* a ranking. Clearly the idea of this procedure is to generate a ranking that contradicts as few pairwise comparisons as possible in the outranking matrix. This is exactly what the Kemeny rule does. While Kemeny's rule always chooses the ranking that is closest to the profile, Tideman's procedure *tries* to find that ranking using *greedy* algorithm. From the computational point of view the Tideman winner is always fast to find while finding the Kemeny winner can be very slow if the number of candidates is large.

The last system considered here is Schulze's method. Let us denote by S_{xy} the maximum strength of all beatpaths from x to y. Using these we generate the 'beatpath tournament matrix' B of the profile P such that the entry $B_{xy} = 1$ if $S_{xy} \geq S_{yx}$; otherwise $B_{xy} = 0$. We say that the ranking R where $R(x) > R(y)$ iff $B_{xy} = 1$ corresponds to this beatpath tournament matrix.

As above we could define the Schulze winning ranking R_{Sc} in profile P with the help of a profile P_{Sc} that is closest to P (with respect to some metric d) and has a beatpath tournament matrix that corresponds to some complete ordering R_{Sc}: Let P_{Sc} be a beatpath tournament matrix that corresponds to complete ordering R_{Sc} and

$$d(P, P_{Sc}) \leq d(P, P')$$

for all profiles P' that have a beatpath tournament matrix that corresponds to some complete ordering. However, Schulze shows that every profile has a beatpath tournament matrix that corresponds to complete ordering. Thus, regardless of the metric, $d(P, P_{Sc}) = 0$ and the winning ranking R_{Sc} is the ordering corresponding to the beatpath tournament matrix of P.

8. Conclusion

The observations made in the preceding are summarized in Table 1. Several entries in the table exhibit goal state-metric combinations that do not correspond to those associated with existing methods. For example, nearly all systems aiming at a Condorcet ordering are purely theoretical, the only exception being Slater's rule. Yet, for a person who finds the Condorcet winning criterion appealing, the Condorcet ordering would probably seem quite a plausible goal set. Depending then on the way this person would like to measure the difference between the observed profile, outranking matrix or tournament matrix, on the one hand, and the profile or matrix representing the Condorcet ordering, he/she might end up with different methods.

As a general conclusion we notice that disagreement may take on several intuitively plausible and yet incompatible meanings. Also most methods that are currently being used can be viewed as devices of looking for the closest consensus, given the expressed opinions of the electorate. Why we have so many different methods turns out, then, to be a natural consequence of our having somewhat different views of what constitutes a consensus and how deviations thereof should be measured.

Table 1. Goal states, metrics and voting systems

Goal state	Unanimous		Condorcet		Beatpath
Metric	winner $\mathbf{W}(x)$	order $U(R)$	winner $\mathbf{C}(x)$	order $\mathbf{Co}(R)$	winner/ order
Inversion d_K	Borda	Kemeny	Dodgson	U 2	Schulze
Manhattan d_L	Borda	Litvak	Dodgson	U 3	Schulze
Inversion for V matrices, d_V	Borda	Kemeny	Condorcet least-reversal	U 1	Schulze
Inversion for T matrices, d_T	–[a]	–[a]	Copeland	Slater	Schulze
Discrete, d_d	Plurality	Plurality	Young	U 4	Schulze

[a] The goal states and the metric are incompatible or their meaning unclear.

Acknowledgements

The authors are indebted to Marlies Ahlert, Matthew Braham, Steven J. Brams, Bernard Grofman, Manfred J. Holler, D. Marc Kilgour, Moshé Machover and an anonymous referee for numerous constructive suggestions.

References

Aleskerov, F. (1999) *Arrovian Aggregation Models*, Kluwer.

Baigent, N. (1987*a*) Preference Proximity and Anonymous Social Choice, *The Quarterly Journal of Economics* 102: 161–169.

Baigent, N. (1987*b*) Metric Rationalization of Social Choice Functions According to Principles of Social Choice, *Mathematical Social Sciences* 13: 59–65.

Barthélemy, J. P. and Monjardet, B. (1981) The Median Procedure in Cluster Analysis and Social Choice Theory, *Mathematical Social Sciences* 1: 235–267.

Black, D. (1958) *Theory of Committees and Elections*, Cambridge University Press.

Brams, S. (1976) One Man, *n* Votes, Module in Applied Mathematics, Mathematical Association of America, Cornell University.

Brams, S. and Fishburn, P. (1983) *Approval Voting*, Birkhäuser.

Doron, G. (1979) The Hare System is Inconsistent, *Political Studies* 27: 283–286.
Farkas, D. and Nitzan, S. (1979) The Borda Rule and Pareto Stability: A Comment, *Econometrica* 47: 1305–1306.
Farquharson, R. (1969) *Theory of Voting*, Yale University Press.
Hare, T. (1865) *Election of Representatives*, Longman.
Hill, I. D. (1988) Some Aspects of Elections – To Fill One Seat or Many, *Journal of the Royal Statistical Society*. Series A 151: 243–275.
Kemeny, J. (1959) Mathematics without Numbers, *Daedalus* 88: 571–591.
Lakeman, E. and Lambert, J. (1955) *Voting in Democracies*, Faber.
McLean, I. and Urken, A. (eds) (1995) *Classics of Social Choice*, University of Michigan Press.
Meskanen, T. and Nurmi, H. (2006) Distance from Consensus: A Theme and Variations, in B. Simeone and F. Pukelsheim (eds) *Mathematics and Democracy: Recent Advances in Voting Systems and Collective Choice*, Springer.
Nanson, E. J. (1882) Methods of Election, *Transactions and Proceedings of the Royal Society of Victoria* XIX: 197–240. Reprinted in I. McLean and A. Urken (1995).
Nitzan, S. (1981) Some Measures of Closeness to Unanimity and Their Implications, *Theory and Decision* 13: 129–138.
Nurmi, H. (2002) Measuring Disagreement in Group Choice Settings, in M. J.Holler et al. (eds), *Power and Fairness*, Jahrbuch für Neue Politische Ökonomie 20, Mohr Siebeck, 313–331.
Nurmi, H. (2004) A Comparison of Some Distance-Based Choice Rules in Ranking Environments, *Theory and Decision* 57: 5–24.
Richelson, J. (1975) A Comparative Analysis of Social Choice Functions, *Behavioral Science* 20: 331–337.
Richelson, J. (1978*a*) A Comparative Analysis of Social Choice Functions II, *Behavioral Science* 23: 38–44.
Richelson, J. (1978*b*) A Comparative Analysis of Social Choice Functions III, *Behavioral Science* 23: 169–178.
Richelson, J. (1979) A Comparative Analysis of Social Choice Functions I,II,III: A Summary, *Behavioral Science* 24: 355.
Richelson, J. (1981) A Comparative Analysis of Social Choice Functions IV, *Behavioral Science* 26: 346–353.
Riker, W. H. (1982) *Liberalism against Populism*, W. H. Freeman.
Schulze, M. (2003) A New Monotonic and Clone-Independent Single-Winner Election Method, *Voting Matters* 17: 9–19.
Slater, P. (1961) Inconsistencies in a Schedule of Paired Comparisons, *Biometrika* 48: 303–312.
Smith, W. D. (2005) Descriptions of Voting Systems, mimeo.
Tideman, N. (1987) Independence of Clones as a Criterion for Voting Rules, *Social Choice and Welfare* 4: 185–206.
Zavist, B. T. and Tideman, N. (1989) Complete Independence of Clones in the Ranked Pairs Rule, *Social Choice and Welfare* 6: 167–173.